

DYNACOMP

SORTIT

1946

SORTIT* Version 2.1
 (C) 1981 by DYNACOMP, Inc.

INTRODUCTION

SORTIT is an 8080 assembly language program that provides efficient sorting of sequential datafiles using the Heapsort algorithm. The program interfaces North Star DOS for disk and user communication. SORTIT is provided in two versions for execution at 2A00H or 2D00H. Both versions handle single/double density, single/double sided drives. The datafile must meet North Star BASIC datafile definitions.

With SORTIT, records may be deleted based on a comparison with a delete key. Records may be sorted in ascending or descending order based on a primary key and optionally, a secondary key. Both keys may be numeric or one to nine character strings with the string in a defined field position or right-justified.

SORTIT requires a detailed definition of the format of the datafile to be sorted. After a user defined description is entered, it may be saved by writing the SORTIT program to disk under a new program name. SORTIT is released with two predefined file descriptions for use with datafiles generated by the DYNACOMP MAILLIST II program, sorting on name or on zip code.

For single drive systems, the source datafile and destination datafile must reside on the same disk. For multiple drive systems each of the two datafiles may exist on any of the drives. SORTIT does NOT write to or update the source datafile.

The CUSTOMIZING SORTIT section of this manual describes the various changes that may be made by the user to configure the program to a particular system.

Memory requirements are determined by the datafile size. See MEMORY REQUIREMENTS section for details.

The following is a brief description of SORTIT operation with emphasis on defining the terminology used in this manual.

The source datafile and sorted destination datafile are identified by one to eight character filenames followed by a comma and the disk drive number. The datafiles are defined to the North Star BASIC requirements. North Star BASIC allows numeric precision of 6, 8, 10, 12 or 14 digits, with 8 - digit precision as standard. The SORTIT customizing table must specify the precision in use.

A SORTIT datafile consists of an optional master record followed by one or more fixed format data records followed by an endmarker.

The master record if present is simply copied to the destination file. It must be defined in the file description as to the number of fields in the record.

The data records are also defined in the file description. They must all be of the same format, that is, the same number of fields and field types.

*Written by Daniel West, North Hollywood, California.

Each data record must contain all defined fields unless an earlier numeric field is present that contains the value of the total number of fields.

Data records may be of variable length in the sense that string fields are of variable length and also in the sense that a specific record may have less than the maximum defined number of fields if a variable field count value is provided. For advanced users, a special fixed length definition is described in a later section of this manual.

SORTIT execution is logically divided into four phases. The initialization phase will display the current file definition and accept a new file definition, then request source and destination filenames. The array build phase reads the source file sequentially and builds a 23 byte entry in system memory for each data record to be sorted. The heapsort phase then sorts these array entries in system memory into the desired order. The copy phase uses the order of the array entries to read each data record from the source file and write it to the destination file.

During the initialization phase, SORTIT displays the current file definition and requests a new file definition and source/destination filenames thru the use of the following prompt lines. The prompt line numbers on the left side of the list are added here for the purpose of easy referencing in later sections of this manual and do not appear with the line on the CRT display. A few of the more obvious prompts are not included in this list.

- 1 - MASTER - # OF FIELDS =
- 2 - DATA RECORD - # OF FIELDS =
- 3 - FIELD # OF VARIABLE FIELD COUNT =
- 4 - VARIABLE # OF FIELDS OFFSET =
- 5 - FIELD # OF DELETE KEY (STRING TYPE ONLY) =
- 6 - LENGTH OF KEY (0 TO 9 ONLY) =
- 7 - OFFSET TO START OF KEY =
- 8 - ENTER DELETE KEY - 1 TO 9 CHARS ;
- 9 - FIELD # OF PRIMARY KEY =
- 10 - FIELD TYPE - NUMERIC OR STRING ? N/S:
- 11 - LENGTH OF KEY (0 TO 9 ONLY) =
- 12 - OFFSET TO START OF KEY =
- 13 - FIELD # OF SECONDARY KEY =
- 14 - FIELD TYPE - NUMERIC OR STRING ? N/S:
- 15 - LENGTH OF KEY (0 TO 9 ONLY) =
- 16 - OFFSET TO START OF KEY =
- 17 - ASCENDING/DESCENDING SORT ? A/D:

If SORTIT is to be used to replace the internal sort of MAILLIST, the reader is advised to study the first three sections of this manual. Users who wish to sort other datafiles should study the entire manual prior to attempting to use SORTIT. The master disk supplied by DYNACOMP should be write protected and used only to copy the program onto a working disk.

The entry of a control-c during the execution of SORTIT has two possible results. If SORTIT has issued a prompt message and is accepting keyboard input, the entry of a control-c anytime up to the entry of a carriage-return will result in an immediate program abort and JUMP to the

DOS reentry address.

Entry of a control-c during the array build phase, sort phase, or copy phase will result in a status message display that will contain a relative processing number.

The record number displayed in the status message is only a relative value to allow the user to verify that processing is continuing. SORTIT will await the entry of a carriage return to continue.

The record numbers displayed at the final status report are true record totals.

CUSTOMIZING SORTIT

The following table lists the various assignments of the first 42 bytes of SORTIT with a comment on the use of each assignment. Modifying these values can result in a customized program. Note that all values are in hex. The values shown are provided in the master copy of the program. The location value is for the 2A00H version and has the same relative offsets at 2D00H for the other version.

| Location | Label | Value | Comment |
|----------|-----------|-----------|------------------------------|
| ----- | ----- | ----- | ----- |
| 2A00 | START | C3 XX XX | JUMP TO PROGRAM START |
| 2A03 | DOUT | C3 0D 20 | TO DOS TERMINAL OUTPUT |
| 2A06 | DINF | C3 10 20 | TO DOS KEYBOARD INPUT |
| 2A09 | CLEAR | C9 00 00 | TO USER SCREEN CLEAR |
| 2A0C | DCOM | C3 22 20 | TO DOS DISK COMMAND |
| 2A0F | DLOOK | C3 1C 20 | TO DOS DIRECTORY SEARCH |
| 2A12 | DOS | C3 28 20 | TO DOS REENTRY |
| 2A15 | CONTC | C3 16 20 | TO DOS STATUS CHECK |
| 2A18 | RUBO | 5F | INPUT DELETE CHARACTER |
| 2A19 | ECHOR | 5F | ECHOED DELETE CHARACTER |
| 2A1A | OPORT | 00 | OUTPUT PORT # |
| 2A1B | IPOINT | 00 | INPUT PORT # |
| 2A1C | MSIZE | 00 00 | MASTER RECORD FIXED SIZE |
| 2A1E | DSIZE | 00 00 | DATA RECORD FIXED SIZE |
| 2A20 | MEMEND | FF 5F | END OF MEMORY |
| 2A22 | PRECISION | 05 | NUMERIC PRECISION |
| 2A23 | SKIPD | 00 | SKIP DISPLAY FLAG |
| 2A24 | SKIPQ | 00 | SKIP FILE DEFINITION REQUEST |
| 2A25 | SKIPF | 00 | SKIP FILENAME KEYIN |
| 2A26 | CASEF | 00 | SKIP CASE CONVERSION |
| 2A27 | SPARE | (3 BYTES) | FUTURE USE |

The Start entry is the execution entry point (not a subroutine call). The next seven entries are all called from within SORTIT. See the North Star DOS descriptions for the required interface requirements. The Clear entry is set to an immediate return. If the user provides the address of a clear screen routine at this point, it will be called during SORTIT initialization.

Rubo and/or Echor may be changed if the user's terminal requires different delete characters.

Msize and Dsize are described in a later section of this manual (FIXED LENGTH RECORDS).

Memend should be set (low order/high order bytes) to the highest available contiguous memory address above 2A00H.

Precision is set to reflect the standard version of North Star BASIC or custom versions for different numeric precisions. Standard 8 - digit precision is set by a value of 5. 6 - digit precision = 4; 10 - digit precision = 6; 12 - digit precision = 7; 14 - digit precision = 8. Non-standard precisions must be matched with the appropriate custom version of BASIC.

Skipd, if set non-zero will result in SORTIT skipping the opening 'Display File Definition' request.

Skipg, if set non-zero will result in SORTIT skipping the 'Define New File' request.

Skipf, if set non-zero will result in SORTIT skipping the source/destination filenames requests.

Casef, if set non-zero will remove the conversion of lowercase letters to uppercase during primary and secondary key comparisons.

MAILLIST II SORTING

The 2A00H version of SORTIT on the master disk has been preset to provide an alphabetic ascending sort on the name field of a MAILLIST file. If executed as is, it will provide the user with a very fast sort. The 2D00H version is preset for an ascending zip code sort. Both versions will delete the appropriately marked records.

You may create a 2A00H version of the zip code sort by first executing the 2D00H version and displaying the file definition. Record the values displayed and cancel the 2D00H version by entering a control-c. Now, prepare a new program area on your working disk for a new copy of SORTIT. Execute the 2A00H version and respond 'Y' to the define file prompt. Respond to each file prompt with the corresponding value you earlier recorded from the 2D00H version. After all prompts have been answered, respond 'Y' to the 'save this file definition' prompt and use the DOS save file (SF) command to save this version under your new program name. You now have a 2A00H version for zip sort.

A similar technique may be used to obtain a 2D00H version of an alpha sort.

[The text on this page is extremely faint and illegible. It appears to be a multi-paragraph document, possibly a letter or a report, but the specific content cannot be discerned.]

The MAILLIST II user may wish to enhance his sort capability with SORTIT versions that provide:

- Descending sorts.
- Delete key processing.
- 1 to 9 character primary key sorts.
- 1 to 9 character secondary key sorts.

A full study of this manual should provide the knowledge needed to enhance the MAILLIST sort in many ways.

MEMORY REQUIREMENTS

SORTIT requires contiguous memory for program, I/O buffers and heapsort array. The available memory for the heapsort array will determine the largest file size that may be sorted. The following table shows the address of the start of the heapsort array depending on SORTIT version and on the density of the two datafiles.

| | 2A00H Version ----- | 2D00H Version ----- |
|--------------------------------|------------------------|------------------------|
| Both datafiles single density: | 4100H | 4400H |
| One datafile double density: | 4200H | 4500H |
| Both datafiles double density: | 4300H | 4600H |

Once the user has obtained the heapsort array start address from the above table, subtract this value from the MEMEND value in the customizing table to obtain the available memory for the heapsort array. Dividing this size (expressed in decimal) by 23 decimal will result in the maximum number of data records that may be sorted. The following is an example.

For a 2D00H version with both datafiles of single density, heapsort array start address is 4400H. If MEMEND is set at 63FFH, then available sort memory is 1FFFFH, which is 8191 decimal. 8191 divided by 23 results in 356 as the number of data records that may be sorted with this memory size.

Note that if a source datafile has data records that will be deleted due to the use of the delete key definition, these records are not included in the above maximum number of sort records.

MASTER RECORD PROCESSING

Some datafiles are built with the first record of the file containing overall file parameters. This initial record (referred to as the master record) may have a different format than the following data records.

If the response to the master number of fields prompt (line # 1) is zero, then SORTIT will not expect the existence of a master record. If the response is 1 to 255, then SORTIT will copy a master record from the source to the destination file as the first record on the file and it will contain the defined number of fields.

DATA RECORD PROCESSING

...the ... of the ...
...the ... of the ...
...the ... of the ...

...the ... of the ...
...the ... of the ...
...the ... of the ...

...the ... of the ...
...the ... of the ...
...the ... of the ...

...the ... of the ...
...the ... of the ...
...the ... of the ...

...the ... of the ...
...the ... of the ...
...the ... of the ...

...the ... of the ...
...the ... of the ...
...the ... of the ...

...the ... of the ...
...the ... of the ...
...the ... of the ...

...the ... of the ...
...the ... of the ...
...the ... of the ...

...the ... of the ...
...the ... of the ...
...the ... of the ...

The data records may contain from 1 to 255 fields (the number specified by response to prompt line # 2) with all records containing the same number of fields (for exception, see below). The fields may be North Star BASIC numeric fields, string fields of 1 to 255 characters or long string fields of 256 or more characters. Each data record must be of the same field format. For example, if the third field of one record is numeric then the third field of all data records must be numeric. The string field of less than 256 characters in one record may correspond to a long string field in other data records, however there are some restrictions on the use of long string fields, described later, that must be considered.

DELETE KEY PROCESSING

A useful feature of SORTIT is the ability to delete records (that is, NOT copy a record to the destination file) that contain a match with a user defined delete key.

The delete key may not be a numeric field and the 1 to 9 character string is entered in the file definition on response to prompt line # 8 with the length of the key in response to prompt line # 6. The area of each data record that is tested for a match is defined by specifying a field number (see prompt line # 5) of other than zero and a key offset value. If the key is left-justified in the field then the offset value should be set to 1 (see prompt line # 7). If the key is right-justified in a string, or right justified in the first 255 positions of a long string field, then the offset value should be set to zero. If the position is fixed within the first 255 bytes of the field, then the offset value should be set to the position of the start of the key area. For example, if the field is 300 bytes and the fifteenth thru the twenty-first characters are the desired delete key, then offset should be set to 15 and length should be set to 7.

Note that the delete key area must always be within the first 255 characters of the field. Each character in the key area is compared to the delete key corresponding position and for deletion the characters must match exactly. Uppercase characters will not match with lowercase characters. There is an exception; if the character in the defined delete key is a question mark (ASCII 3FH), then any character in the corresponding data record area will be considered a match. Note that a defined delete key consisting of all question marks would result in deleting every record that contained a field of sufficient length to have at least all the specified delete key length.

PRIMARY AND SECONDARY KEY PROCESSING

The primary and secondary keys are defined in the same way and have the same characteristics. A primary key field number must always be specified (see prompt line # 9) but a secondary key need not be specified (response to prompt line # 13 of zero). The rest of this section will refer to either key.

The key field may be numeric or string. If numeric, an ascending sort will select the most negative values, then zero, then positive values using 8 digit precision or optionally one of the non-standard precision values. Descending sorts will reverse the above order.

The first part of the report deals with the general situation of the country. It is a very interesting and informative study of the country's development. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is a valuable contribution to the study of the country's development.

The second part of the report deals with the economic situation of the country. It is a very interesting and informative study of the country's economic development. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is a valuable contribution to the study of the country's economic development.

The third part of the report deals with the social situation of the country. It is a very interesting and informative study of the country's social development. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is a valuable contribution to the study of the country's social development.

The fourth part of the report deals with the political situation of the country. It is a very interesting and informative study of the country's political development. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is a valuable contribution to the study of the country's political development.

The fifth part of the report deals with the cultural situation of the country. It is a very interesting and informative study of the country's cultural development. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is a valuable contribution to the study of the country's cultural development.

If the key field is a string type the definition and record area useage is the same as described for the delete key except that all lowercase letters are converted to uppercase letters for the comparison testing and there is no use of the question mark delimiter. The customizing section describes how to disable the lowercase to uppercase conversion, if desired.

If the record length is not sufficient to contain the full number of characters defined by the key length, then the reduced number of characters will be used. If the field is so short as to contain no characters to use for sorting, this record will appear at the front of the sorted file on ascending sorts, or at the end of the sorted file on descending sorts.

VARIABLE FIELD COUNT PROCESSING

Some programs that build datafiles use the technique of providing a field counter value in one of the first fields of each data record. The data records are defined to contain a certain minimum number of fields and a certain maximum number of fields. The field counter value in each data record defines the total number of fields present for that particular record.

SORTIT will process this type of datafile by allowing the definition of a 'variable field count' and a variable offset value. The maximum number of fields in a data record is specified in the response to the number of fields prompt (line # 2). The location of the numeric field to be used for the variable field counter is used in response to the 'field number of variable field count' (prompt line # 3). If, for example, the second field of each record is a numeric field containing a value equal to the number of remaining fields in the record, then response to prompt line # 3 would be 2. The actual value contained in this second field would be zero if no additional fields were in this particular record. SORTIT requires an offset value to be defined that when added to the value contained in the variable count field will result in the total number of fields present. For the above example, the response to the offset prompt (line # 4) should be 2. If the data record contains four fields, the contents of the second field (2) will be added to the offset (2) to obtain the total number of fields (4).

As another example, the MAILLIST datafile contains variable field size records from 5 to 7 fields, with a variable field count value from 3 to 5 in the first field of each data record. Response to the number of fields prompt (line # 2) is 7; to the field number of the variable count prompt (line # 3) is 1; and to the field count offset prompt (line # 4) is 2.

The optional master record does not use this variable field count, the record must contain the exact number of fields specified in response to prompt line # 1.

The variable field count field must be of numeric type.

If the file is of this variable count type and any of the key definitions are specified as right-justified, then the key will be obtained from its defined field number -OR- from the last field in the record if there are less fields in the record than the defined field number. In other words,

1. The first part of the report deals with the general situation of the country. It is a very interesting and informative study of the country's development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

2. The second part of the report deals with the economic situation of the country. It is a very interesting and informative study of the country's economic development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

3. The third part of the report deals with the social situation of the country. It is a very interesting and informative study of the country's social development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

4. The fourth part of the report deals with the political situation of the country. It is a very interesting and informative study of the country's political development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

5. The fifth part of the report deals with the cultural situation of the country. It is a very interesting and informative study of the country's cultural development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

6. The sixth part of the report deals with the environmental situation of the country. It is a very interesting and informative study of the country's environmental development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

7. The seventh part of the report deals with the international situation of the country. It is a very interesting and informative study of the country's international development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

8. The eighth part of the report deals with the future of the country. It is a very interesting and informative study of the country's future development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

9. The ninth part of the report deals with the conclusion of the study. It is a very interesting and informative study of the country's development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

10. The tenth part of the report deals with the appendix. It is a very interesting and informative study of the country's development and progress. The author has done a great deal of research and has gathered a wealth of material. The report is well written and is easy to read. It is a valuable contribution to the knowledge of the country and its people.

the key area is right-justified in the field -AND- in the record, if the record is sufficiently short.

FIXED LENGTH RECORDS

The previous sections have described the data records as if they were part of a true sequential file, with each record immediately following another.

Some access methods place data records at fixed length blocks in the datafile. If a particular record is smaller than the assigned space, an unused number of bytes occur between the record and the next record in the file.

For this type of 'random access' file technique, SORTIT may be customized to recognize and advance to the next record by stepping past the unused bytes.

The user must set the MSIZE value (low order byte first) in the customizing table to the total fixed length for the master record, if present. For the data records the user must set the DSIZE value (low order byte first) in the customizing table. All other aspects of the data records continue to be as described throughout this manual including the variable field count capabilities. Of course, the record must never exceed the fixed length value. The datafile endmarker (01H) must occur on a fixed record boundary and not immediately following the last byte of the last field.

ERROR PROCESSING

During file specification, the entry of an undefined or out of range value will result in the immediate repeat of the same prompt line. For example, entry of a 12 to a length prompt (key lengths are 1 to 9) will result in the same prompt being repeated.

If both datafiles are on the same drive at the same location (maybe the same filename), SORTIT will repeat the filenames prompts.

If a file is not present on the specified drive, SORTIT will output the message '-filename- NOT FOUND ON DRIVE, RETRY? Y/N'. If you then properly load the drive and respond 'Y', SORTIT will continue. Any other response will cause an abort and reentry to DOS.

If, after the initial check and during later processing, a file is not found on the specified drive the message 'DISK FILE NO LONGER AVAILABLE' is displayed and SORTIT will reenter DOS.

If during processing, a field is found of a different type than expected, SORTIT will display 'UNKNOWN RECORD TYPE DETECTED' and will reenter DOS.

If the datafile contains a greater number of records than may be processed in the Heapsort array memory area, SORTIT will display 'SORRY, INSUFFICIENT MEMORY' and will reenter DOS.

If SORTIT does not recognize an endmarker in the source datafile, or cannot fit all of the records in the destination datafile, the 'DISK FILE SIZE PROBLEM' message will be displayed and DOS will be reentered.

The final status report displayed by SORTIT indicates the total records

The second part of the report is a detailed account of the work done during the year.

The third part of the report is a summary of the work done during the year.

The fourth part of the report is a detailed account of the work done during the year.

The fifth part of the report is a summary of the work done during the year.

The sixth part of the report is a detailed account of the work done during the year.

The seventh part of the report is a summary of the work done during the year.

The eighth part of the report is a detailed account of the work done during the year.

The ninth part of the report is a summary of the work done during the year.

The tenth part of the report is a detailed account of the work done during the year.

The eleventh part of the report is a summary of the work done during the year.

The twelfth part of the report is a detailed account of the work done during the year.

copied (including the master record, if present); the number of records deleted due to match on the delete key; the number of bad primaries and number of bad secondaries. A bad primary/secondary occurs when the field type does not match the specified type. These records will be moved to the sorted file and the total records copied reflects this move. They will be at the front of the file on ascending sorts and at the back on descending sorts.

